

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
11 October 2001 (11.10.2001)

PCT

(10) International Publication Number
WO 01/75067 A2

(51) International Patent Classification⁷: C12N

TANG, Y., Tom [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US).

(21) International Application Number: PCT/US01/08631

(22) International Filing Date: 30 March 2001 (30.03.2001)

(74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Cohn, Ferris, Glovsky and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

09/540,217	31 March 2000 (31.03.2000)	US
09/649,167	23 August 2000 (23.08.2000)	US

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	09/540,217 (CIP)
Filed on	31 March 2000 (31.03.2000)
US	09/649,167 (CIP)
Filed on	23 August 2000 (23.08.2000)

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (*for all designated States except US*): HYSEQ, INC. [US/US]; 670 Almanor Avenue, Sunnyvale, CA 94086 (US).

Published:

- without international search report and to be republished upon receipt of that report
- with sequence listing part of description published separately in electronic form and available upon request from the International Bureau

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): DRMANAC, Rodoje, T. [YU/US]; 850 East Greenwich Place, Palo Alto, CA 94303 (US). LIU, Chenghua [CN/US]; 1125 Ranchero Way, Apt. #14, San Jose, CA 95117 (US).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and uses thereof.



WO 01/75067 A2

WHAT IS CLAIMED IS:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-30368, a mature protein coding portion of SEQ ID NO: 1-30368, an active domain of SEQ ID NO: 1-30368, and complementary sequences thereof.
- 5 2. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide hybridizes to the polynucleotide of claim 1 under stringent hybridization conditions.
- 10 3. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide has greater than about 90% sequence identity with the polynucleotide of claim 1.
4. The polynucleotide of claim 1 wherein said polynucleotide is DNA.
- 15 5. An isolated polynucleotide of claim 1 wherein said polynucleotide comprises the complementary sequences.
6. A vector comprising the polynucleotide of claim 1.
- 20 7. An expression vector comprising the polynucleotide of claim 1.
8. A host cell genetically engineered to comprise the polynucleotide of claim 1.
9. A host cell genetically engineered to comprise the polynucleotide of claim 1 operatively
25 associated with a regulatory sequence that modulates expression of the polynucleotide in the host cell.
10. An isolated polypeptide, wherein the polypeptide is selected from the group consisting of:
 - (a) a polypeptide encoded by any one of the polynucleotides of claim 1; and
 - 30 (b) a polypeptide encoded by a polynucleotide hybridizing under stringent conditions with any one of SEQ ID NO: 1-30368.
11. A composition comprising the polypeptide of claim 10 and a carrier.
- 35 12. An antibody directed against the polypeptide of claim 10.

of these novel nucleic acids. Such DNA sequences include those which are capable of hybridizing to the appropriate novel nucleic acid sequence under stringent conditions.

Polynucleotides encoding preferred polypeptide truncations of the invention can be used to generate polynucleotides encoding chimeric or fusion proteins comprising one or more domains of the invention and heterologous protein sequences.

The polynucleotides of the invention additionally include the complement of any of the polynucleotides recited above. The polynucleotide can be DNA (genomic, cDNA, amplified, or synthetic) or RNA. Methods and algorithms for obtaining such polynucleotides are well known to those of skill in the art and can include, for example, methods for determining hybridization conditions that can routinely isolate polynucleotides of the desired sequence identities.

In accordance with the invention, polynucleotide sequences comprising the mature protein coding sequences corresponding to any one of SEQ ID NO: 1-30368, or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of that nucleic acid, or a functional equivalent thereof, in appropriate host cells. Also included are the cDNA inserts of any of the clones identified herein.

A polynucleotide according to the invention can be joined to any of a variety of other nucleotide sequences by well-established recombinant DNA techniques (see Sambrook J et al. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, NY). Useful nucleotide sequences for joining to polynucleotides include an assortment of vectors, e.g., plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a polynucleotide of the invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism, convenient restriction endonuclease sites, and a selectable marker for the host cell. Vectors according to the invention include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

The present invention further provides recombinant constructs comprising a nucleic acid having any of the nucleotide sequences of SEQ ID NO: 1-30368 or a fragment thereof or any other polynucleotides of the invention. In one embodiment, the recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a nucleic acid having any of the nucleotide sequences of SEQ ID NO: 1-30368 or a fragment thereof is inserted, in a forward or reverse orientation. In the case of a vector comprising one of the ORFs of the present invention, the vector may further comprise regulatory sequences, including for example, a promoter, operably linked to the ORF. Large numbers of suitable vectors and

promoters are known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of example. Bacterial: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pDR540, pRIT5 (Pharmacia).

5 Eukaryotic: pWLneo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al., *Nucleic Acids Res.* 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many
10 suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, *Methods in Enzymology* 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed
15 (transfected) with the ligated polynucleotide/expression control sequence.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine
20 kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art. Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, *e.g.*, the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly-expressed gene to direct
25 transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), a-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the
30 periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an amino terminal identification peptide imparting desired characteristics, *e.g.*, stabilization or simplified purification of expressed recombinant product. Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination
35 signals in operable reading phase with a functional promoter. The vector will comprise one or

more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced or derepressed by appropriate means (*e.g.*, temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Polynucleotides of the invention can also be used to induce immune responses. For example, as described in Fan et al., *Nat. Biotech.* 17:870-872 (1999), incorporated herein by reference, nucleic acid sequences encoding a polypeptide may be used to generate antibodies against the encoded polypeptide following topical administration of naked plasmid DNA or following injection, and preferably intramuscular injection of the DNA. The nucleic acid sequences are preferably inserted in a recombinant expression vector and may be in the form of naked DNA.

4.3 ANTISENSE

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 1-30368, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a protein of any of SEQ ID

combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, *e.g.*, RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn *et al.* (1996) *Nucl Acids Res* 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag *et al.* (1989) *Nucl Acid Res* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn *et al.* (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen *et al.* (1975) *Bioorg Med Chem Lett* 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.*, 1989, *Proc. Natl. Acad. Sci. U.S.A.* 86:6553-6556; Lemaitre *et al.*, 1987, *Proc. Natl. Acad. Sci.* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, *e.g.*, Krol *et al.*, 1988, *BioTechniques* 6:958-976) or intercalating agents (see, *e.g.*, Zon, 1988, *Pharm. Res.* 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

4.5 HOSTS

The present invention further provides host cells genetically engineered to contain the polynucleotides of the invention. For example, such host cells may contain nucleic acids of the invention introduced into the host cell using known transformation, transfection or infection methods. The present invention still further provides host cells genetically engineered to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell.

Knowledge of nucleic acid sequences allows for modification of cells to permit, or increase, expression of endogenous polypeptide. Cells can be modified (*e.g.*, by homologous

recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the polypeptide at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the encoding sequences. See, for example, PCT International Publication No. WO94/12650, PCT International Publication No. WO92/20808, and PCT International Publication No. WO91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (*e.g.*, *ada*, *dhfr*, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the recombinant construct into the host cell can be effected by calcium phosphate transfection, DEAE, dextran mediated transfection, or electroporation (Davis, L. et al., *Basic Methods in Molecular Biology* (1986)). The host cells containing one of the polynucleotides of the invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

Any host/vector system can be used to express one or more of the ORFs of the present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, Cv-1 cell, COS cells, 293 cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular polypeptide or protein or which expresses the polypeptide or protein at low natural level. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., in *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor, New York (1989), the disclosure of which is hereby incorporated by reference.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, *Cell* 23:175 (1981). Other cell lines capable of expressing a compatible vector are, for example, the C127, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3

cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Recombinant polypeptides and proteins produced in bacterial culture are usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or insects or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli*, *Bacillus subtilis*, *Salmonella typhimurium*, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequence include polyadenylation signals, mRNA stability elements, splice

sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, *e.g.*, inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the host cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (*gpt*) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.6 POLYPEPTIDES OF THE INVENTION

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising: the amino acid sequences set forth as any one of SEQ ID NO: 30369-60736 or an amino acid sequence encoded by any one of the nucleotide sequences SEQ ID NO: 1-30368 or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides preferably with biological or immunological activity that are encoded by: (a) a polynucleotide having any one of the nucleotide sequences set forth in SEQ ID NO: 1-30368 or

hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other hematopoietic factors, protein or other active ingredient of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein or other active ingredient of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

4.12.1 ROUTES OF ADMINISTRATION

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections. Administration of protein or other active ingredient of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into a arthritic joints or in fibrotic tissue, often in a depot or sustained release formulation. In order to prevent the scarring process frequently occurring as complication of glaucoma surgery, the compounds may be administered topically, for example, as eye drops. Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with a specific antibody, targeting, for example, arthritic or fibrotic tissue. The liposomes will be targeted to and taken up selectively by the afflicted tissue.

The polypeptides of the invention are administered by any route that delivers an effective dosage to the desired site of action. The determination of a suitable route of administration and an effective dosage for a particular indication is within the level of skill in the art. Preferably for wound treatment, one administers the therapeutic compound directly to the site. Suitable dosage ranges for the polypeptides of the invention can be extrapolated from these dosages or from similar studies in appropriate animal models. Dosages can then be adjusted as necessary by the clinician to provide maximal therapeutic benefit.

4.12.2 COMPOSITIONS/FORMULATIONS

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in a conventional manner using one or more physiologically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. These pharmaceutical compositions may be manufactured in a manner that is itself known, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes. Proper formulation is dependent upon the route of administration chosen. When a therapeutically effective amount of protein or other active ingredient of the present invention is administered orally, protein or other active ingredient of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein or other active ingredient of the present invention, and preferably from about 25 to 90% protein or other active ingredient of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein or other active ingredient of the present invention, and preferably from about 1 to 50% protein or other active ingredient of the present invention.

When a therapeutically effective amount of protein or other active ingredient of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein or other active ingredient of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein or other active ingredient solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein or other active ingredient of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate

<210> 28100
<211> 5027
<212> DNA
<213> Homo sapiens

<220>
<221> SIMILAR
<222> (1906)..(2340)
<223> 100% homologous to Homo sapiens dJ686C3.3 (novel
gene), accession number AL049712, Smith-Waterman Score=778.

<400> 28100
atgggggtttc accatgtttg ccacgctggt ctggaacacc tgacctcagg tgacaggctg 60
ggaaggagat cctcaagcaa gcgggctctc aaagccgagg ggaccccagg caggcgcgga 120
gctcagcgaa gccagaagga ggcgcgccgg ggcagcccaa gcccggggtc tccccggagg 180
aagcaaacag ggcgcaggag acacagagaa gagctggggg agcaggagcg gggcgaggca 240
gagaggacct gcgagggcag gagaaagcgc gacgagaggg cctccttcca ggagcggaca 300
gcagcccaa agagggaaaa ggagattccg aggaggaggg agaagtcgaa gcggcagaag 360
aaaccagggt catcctcctt ggctccagt gcctctggtg gggagtcctt gtccgaggag 420
gaactggccc ggactctgga gcaggtgga gaaaaaaga agctcattgc caccatgcgg 480
agcaagccct ggcccatggc gaagaagctg acagagctca gggaggccca ggaatttgtg 540
gagaagtatg aaggcgcctt gggaaagggg aaaggcaagc aactatatgc ctacaagatg 600
ctgatggcca agaaatgggt caaattttaag agagactttg ataatttcaa gactcaatgt 660
atcccttggg aaatgaagat caaggacatt gaaagtcact ttggttcttc agtggcatcg 720
tatttcatct ttctccgatg gatgtatgga gttaaccttg tcttttttgg cttaatatgt 780
ggtctagtca taatcccaga ggatgtctac gtgatccctg aggaaccctc agttatgctg 840
caggagctgg ctggcaaggc cccactggat gacaagttaa tgtacttctc ttccaacact 900
ggatcctacg gtctagggtt tgaccaaggc tacaattatc ttgaggctga actgaagaag 960
atccgcttcc aagctcactc acatggctac tggcagatct cagaagatac aatttcaagc 1020
ttactcacgt ggcttctggc aggcctcaga aaatctgctt ccaaccttac ttatgtggct 1080
gttcccaagg ttccagggcta tatcaagtac tctgcaactc tctatggcta ctacaacaac 1140
cagaggacca tcgggtggct gaggtaccgg ctgcctatgg ctactttat ggtgggggtc 1200
agcgtgttcg gctacagcct gattattgtc attcgatcga tggccagcaa tacccaagga 1260
agcacaggcg aaggggagag tgacaacttc acattcagct tcaagatgtt caccagctgg 1320
gactacctga tcgggaattc agagacagct gataacaaat atgcatccat caccaccagc 1380
ttcaaggaat caatagtgga tgaacaagag agtaacaaag aagaaaatat ccatctgaca 1440
agatttcttc gtgtcctggc caactttctc atcatctgct gtttgtgtgg aagtgggtac 1500
ctcatttact ttgtggttaa gcgatctcag caattctcca aaatgcagaa tgtcagctgg 1560
tatgaaagga atgaggtaga gatcgtgatg tccctgcttg gaatgttttg tccccctctg 1620
tttgaaacca tcgctgccct ggagaattac caccacgca ctggactgaa gtggcagctg 1680
ggacgcatct ttgcaactct cctggggaac ctctacacat ttctcttggc cctgatggat 1740
gacgtccacc tcaagcttgc taatgaagag acaataaaga acatcactca ctggactctg 1800
tttaactatt acaactcttc tgggttggaa gagagtgtcc cccgaccacc cctgcacct 1860
gcagatgtgc cccgggggtc ttgctgggag acagctgtgg gcattgaatt catgaggctg 1920
acggtgtctg acatgtctgg aacgtacatc accatcctgc tgggggactt cctacgggct 1980
tgttttctgc ggttcatgaa ctactgctgg tgctgggact tggaggctgg atttccctca 2040
tatgctgagt ttgatattag tggaaatgtg ctgggtttga tcttcaacca aggaatgatc 2100
tggatgggct ccttctatgc tccaggcctg gtgggcatta atgtgctgcg cctgctgacc 2160
tccatgtact tccagtgtcg ggcggtgatg agcagcaacg taccatga acgcgtgttc 2220
aaagcctccc gatccaaca cttctacatg ggctcctgc tgctggtgct ctctctcagc 2280
ctcctgcggg tggcctacac catcatgtcc ctcccacct cctttgactg cgggcggtc 2340
agtgggaaaa acagaatgta cgatgtcctc caagagacca ttgaaaacga tttcccaacc 2400
ttcctgggca agatctttgc tttcctcgcc aatccaggcc tgatcatccc agccatcctg 2460
ctgatgttct tggccattta ctacctgaac tcagtttcca aaagccttcc cagactaat 2520
gcccagctga ggaagaaaa ccaagtgtc cgtgaagtgg agaagagtca caaatctgta 2580
aaaggcaaa ccacagccag agattcagag gacacaccta aaagcagctc caaaaatgcc 2640

accagctcc	aactaccaa	ggaagagacc	actcctccct	ctgccagcca	aagccaggcc	2700
atggacaaga	aggcgaggg	ccctgggacc	tccaattctg	ccagcaggac	cacactgcct	2760
gcctctggac	accttcctat	atctcgcccc	cctggaatcg	gaccagattc	tggccacgcc	2820
ccatctcaga	ctcatccgtg	gagacagggc	ctgggcctgg	gcctgggcct	gcgcctgcgc	2880
ctgcgcctgc	cctgggaacg	ggttccggca	gacgtgagg	ttgcgttgac	gctcgcgccc	2940
cggtccccgt	tccaggtgct	gttgacgtg	tctgtttgag	cacgcaggtc	ggctacacgc	3000
atgctggcgc	tgaaggaag	tggaggagat	cagtctgctg	cagccgcagg	tgggaggagt	3060
ccgtgctcaa	cctggggcaa	attccacagc	atcagttcgt	ctggtggcct	tttgtccctt	3120
tgcctcatcc	caggttgccct	tggaaaatag	ccaacgccgt	gtctgaaagg	ggttgttcat	3180
gaggacctcc	gcctgctctt	ggaagaccca	cctgccgtcc	aaaaagaaga	aaagtactct	3240
tggggagttg	ggggatccca	aagattggtg	ccgcaataca	ggaaggagtt	aggggtacaac	3300
ttgccagact	ggaaggaatc	atagctgaga	ttcctgcgag	gagttcgtct	gcactttcca	3360
caaatctggg	ttgaaagggg	ctgaaccgat	tctgtcagct	ttgtaaagca	caagctgggg	3420
cttggggacac	agctattccc	ggtgccaaaa	gttaagttta	atgtgaaccc	gggtggacaa	3480
tatgatcatc	cagtccatta	gcctccctgg	accaagctgg	ataaaggaca	tcaatacctt	3540
ctctaatagcc	gtgtcaaggg	agtggtagcg	ggtaatcact	ttccggagct	tggttgaaga	3600
tcatcaaccg	acaatgccac	atactgccgt	cttgcccagt	tttattggaa	accgaaggag	3660
acttgaatga	ggacaagcat	ggagaagctg	gaggagctga	caatggatgg	ggcccaaggc	3720
taaggctatt	ctggatgcct	cacggtcctc	catgggcattg	gacatatctg	ccactgactt	3780
gataaacatc	gagagcttct	ccagtcgtgt	ggtgtcttta	tctgaatacc	ggccagagcc	3840
tacacactta	cctgcgctcc	aagatgagcc	aagttagcccc	cagcctgtca	gccctaattg	3900
gggaagcggg	aggggtgcacg	tcttcacgcg	acatgcttgg	cagccttcac	caacctgggc	3960
caagtattcc	agcattccac	agtgcgaagat	cttgggggct	gaaaaggccc	tgttcagagc	4020
cctgaagaca	aggggttaaca	ccccaaaatt	atgggaactcc	attttcccac	tccaccttc	4080
attgggccga	gcagctgccc	aagaactaaa	ggcccgcatc	tccccgatac	cctggcaaac	4140
aaatgcagta	ttgccctcac	gaatccgatt	gcttctctgg	ctttgtcacc	cacacacatc	4200
cagaggtgcc	cacgaagtgt	attcggggag	aagcttcgag	aacaaagttg	aagagcgact	4260
gtccttctat	gagactggag	agataccacg	gaaagaatct	ggatgtcatg	aaggaagcaa	4320
tggttcaagc	aagaggcaga	ggaagcggct	gctgaggatt	actaggggaag	ctggagaaac	4380
aggagaagaa	acgcttaaa	aaggaaaaga	accgggctgg	cttgcacttt	gcccctcgcg	4440
tcttcaggaa	accagcagtt	agttcctcca	gaggagttgt	tgaggaagac	gagtgaaaac	4500
ccccaaaaag	gaaggaaaaa	gccaaaagcc	ccccaggagg	tttccttcag	ggagaatggg	4560
aattgggaag	acccattctt	atcttctttt	ctcccaaaac	cccaaggaaa	aagaaatctt	4620
tttcccaagg	aggagttgat	gagtttagcg	tcttgaaga	gacccgctgg	cagcacccag	4680
ttattcccaa	gaggaagaag	tctacaccca	aggaggaaac	agttaatgac	cccttgggaag	4740
gcaggccaca	aaaagtggct	ccaagaaaaa	ggaggaaatt	ctccaaagga	ggagccggct	4800
aagcagtggg	cctgaagagc	cggttggcaa	gagcagctcc	aagaagaaga	aaaagttcca	4860
taaagcatcc	caggaagatt	agaatgcaaa	tggacattct	ctgggaggtg	gggcatacca	4920
tagcccaagg	tgctcatttc	ccacctgtg	cccggtgtcc	ccaataaaaa	caaattcaca	4980
agaaaaaaaa	aaaaaaaaaa	aaaattcctg	aggccgcaag	ggaattc		5027

AAS92296

ID AAS92296 standard; cDNA; 5027 BP.

XX

AC AAS92296;

XX

DT 13-FEB-2002 (first entry)

XX

DE DNA encoding novel human diagnostic protein #28100.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.

XX

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Drmanac RT, Liu C, Tang YT;

XX

DR WPI; 2001-639362/73.

DR P-PSDB; ABG28109.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX

PS Claim 1; SEQ ID NO 28100; 103pp; English.

XX

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at

Qy	813	TTTTGGCTTAATATTTGGTCTAGTCATAATCCCAGAGGTACT-----	854
Db	765	TTTTGGCTTAATATTTGGTCTAGTCATAATCCCAGAGGATGTCTACGTGATCCCTGAGGA	824
Qy	855	-----GATGGGCATGCCCTATGGGAGTATTCCCAGAAAGA	889
Db	825	ACCCTCAGTTATGCTGCAGGAGCTGGCTGGCAAGGCCCCACTGGATGACAAGTTTATGTA	884
Qy	890	CAGTGCCTCGGGCTGAGGAAGAAAAGGCCATGGATTTTTCT-----	930
Db	885	CTTCTCTTCCAACACTGGATCCTACGGTCTAGGTTTTGACCAAGGCTACAATTATCTTGA	944
Qy	931	-----	930
Db	945	GGCTGAACTGAAGAAGATCCGCTTCCAAGCTCACTCACATGGCTACTGGCAGATCTCAGA	1004
Qy	931	-----	930
Db	1005	AGATACAATTTCAAGCTTACTCACGTGGCTTCTGGCAGGCCTCAGAAAATCTGCTTCCAA	1064
Qy	931	-----GTCCTTTGGGATTTTGAGGGCTATATCAAGTACTCTGCACTCTTCTA	977
Db	1065	CCTTACTTATGTGGCTGTTCCCAAGGTTGAGGGCTATATCAAGTACTCTGCACTCTTCTA	1124
Qy	978	TGGCTACTACAACAACCAGAGGACCATCGGGTGGCTGAGGTACCGGCTGCCTATGGCTTA	1037
Db	1125	TGGCTACTACAACAACCAGAGGACCATCGGGTGGCTGAGGTACCGGCTGCCTATGGCTTA	1184
Qy	1038	CTTTATGGTGGGGGTGAGCGTGTTCGGCTACAGCCTGATTATTGTCATTTCGATCGATGGC	1097
Db	1185	CTTTATGGTGGGGGTGAGCGTGTTCGGCTACAGCCTGATTATTGTCATTTCGATCGATGGC	1244
Qy	1098	CAGCAATACCCAAGGAAGCACAGGCGAAGGGGAGAGTGACAACCTTCACATTTCAGCTTCAA	1157
Db	1245	CAGCAATACCCAAGGAAGCACAGGCGAAGGGGAGAGTGACAACCTTCACATTTCAGCTTCAA	1304
Qy	1158	GATGTTTACCAGCTGGGACTACCTGATCGGGAATTTCAGAGACAGCTGATAACAAATATGC	1217
Db	1305	GATGTTTACCAGCTGGGACTACCTGATCGGGAATTTCAGAGACAGCTGATAACAAATATGC	1364
Qy	1218	ATCCATCACCACCAGCTTCAAGGAATCAATAGTGGATGAACAAGAGAGTAACAAAGAAGA	1277
Db	1365	ATCCATCACCACCAGCTTCAAGGAATCAATAGTGGATGAACAAGAGAGTAACAAAGAAGA	1424
Qy	1278	AAATATCCATCTGACAAGATTTCTTCGTGTCTGGCCAACTTTCTCATCATCTGCTGTTT	1337
Db	1425	AAATATCCATCTGACAAGATTTCTTCGTGTCTGGCCAACTTTCTCATCATCTGCTGTTT	1484
Qy	1338	GTGTGGAAGTGGGTACCTCATTTACTTTGTGGTTAAGCGATCTCAGCAATTCTCCAAAAT	1397
Db	1485	GTGTGGAAGTGGGTACCTCATTTACTTTGTGGTTAAGCGATCTCAGCAATTCTCCAAAAT	1544
Qy	1398	GCAGAAATGTCAGCTGGTATGAAAGGAATGAGGTAGAGATCGTGATGTCCCTGCTTGGAAAT	1457
Db	1545	GCAGAAATGTCAGCTGGTATGAAAGGAATGAGGTAGAGATCGTGATGTCCCTGCTTGGAAAT	1604

Qy	1458	GTTTTGTCCCCCTCTGTTTGAAACCATCGCTGCCCTGGAGAATTACCACCCACGCACTGG	1517
Db	1605	GTTTTGTCCCCCTCTGTTTGAAACCATCGCTGCCCTGGAGAATTACCACCCACGCACTGG	1664
Qy	1518	ACTGAAGTGGCAGCTGGGACGCATCTTTGCACTCTTCCTGGGGAACCTCTACACATTTCT	1577
Db	1665	ACTGAAGTGGCAGCTGGGACGCATCTTTGCACTCTTCCTGGGGAACCTCTACACATTTCT	1724
Qy	1578	CTTGGCCCTGATGGATGACGTCCACCTCAAGCTTGCTAATGAAGAGACAATAAAGAACAT	1637
Db	1725	CTTGGCCCTGATGGATGACGTCCACCTCAAGCTTGCTAATGAAGAGACAATAAAGAACAT	1784
Qy	1638	CACTCACTGGACTCTGTTTAACTATTACAACCTCTTCTGGTTGGAACGAGAGTGTCCCCCG	1697
Db	1785	CACTCACTGGACTCTGTTTAACTATTACAACCTCTTCTGGTTGGAACGAGAGTGTCCCCCG	1844
Qy	1698	ACCACCCCTGCACCCCTGCAGATGTGCCCCGGGGTTCTTGCTGGGAGACAGCTGTGGGCAT	1757
Db	1845	ACCACCCCTGCACCCCTGCAGATGTGCCCCGGGGTTCTTGCTGGGAGACAGCTGTGGGCAT	1904
Qy	1758	TGAATTCATGAGGCTGACGGTGTCTGACATGCTGGTAACGTACATCACCATCCTGCTGGG	1817
Db	1905	TGAATTCATGAGGCTGACGGTGTCTGACATGCTGGTAACGTACATCACCATCCTGCTGGG	1964
Qy	1818	GGACTTCCTACGGGCTTGTTTTGTGCGGTTTCATGAACCTACTGCTGGTGCTGGGACTTGGA	1877
Db	1965	GGACTTCCTACGGGCTTGTTTTGTGCGGTTTCATGAACCTACTGCTGGTGCTGGGACTTGGA	2024
Qy	1878	GGCTGGATTTCTTCATATGCTGAGTTTGATATTAGTGGAAATGTGCTGGGTTTGATCTT	1937
Db	2025	GGCTGGATTTCTTCATATGCTGAGTTTGATATTAGTGGAAATGTGCTGGGTTTGATCTT	2084
Qy	1938	CAACCAAGGAATGATCTGGATGGGCTCCTTCTATGCTCCAGGCCTGGTGGGCATTAATGT	1997
Db	2085	CAACCAAGGAATGATCTGGATGGGCTCCTTCTATGCTCCAGGCCTGGTGGGCATTAATGT	2144
Qy	1998	GCTGCGCCTGCTGACCTCCATGTACTTCCAGTGCTGGGCGGTGATGAGCAGCAACGTACC	2057
Db	2145	GCTGCGCCTGCTGACCTCCATGTACTTCCAGTGCTGGGCGGTGATGAGCAGCAACGTACC	2204
Qy	2058	CCATGAACGCGTGTTCAAAGCCTCCCGATCCAACAACCTTCTACATGGGCCTCCTGCTGCT	2117
Db	2205	CCATGAACGCGTGTTCAAAGCCTCCCGATCCAACAACCTTCTACATGGGCCTCCTGCTGCT	2264
Qy	2118	GGTGCTCTTCCTCAGCCTCCTGCCGGTGGCCTACACCATCATGTCCCTCCCACCCTCCTT	2177
Db	2265	GGTGCTCTTCCTCAGCCTCCTGCCGGTGGCCTACACCATCATGTCCCTCCCACCCTCCTT	2324
Qy	2178	TGACTGCGGGCCGTTTCAGTGGGAAAAACAGAATGTACGATGTCCTCCAAGAGACCATTGA	2237
Db	2325	TGACTGCGGGCCGTTTCAGTGGGAAAAACAGAATGTACGATGTCCTCCAAGAGACCATTGA	2384
Qy	2238	AAACGATTTCCCAACCTTCCTGGGCAAGATCTTTGCTTTCCTCGCCAATCCAGGCCTGAT	2297
Db	2385	AAACGATTTCCCAACCTTCCTGGGCAAGATCTTTGCTTTCCTCGCCAATCCAGGCCTGAT	2444

Qy	2298	CATCCCAGCCATCCTGCTGATGTTCTTGGCCATTTACTACCTGAACTCAGTTTCCAAAAG	2357
Db	2445	CATCCCAGCCATCCTGCTGATGTTCTTGGCCATTTACTACCTGAACTCAGTTTCCAAAAG	2504
Qy	2358	CCTTTCCCGAGCTAATGCCCAGCTGAGGAAGAAAATCCAAGTGCTCCGTGAAGTTGAGAA	2417
Db	2505	CCTTTCCCGAGCTAATGCCCAGCTGAGGAAGAAAATCCAAGTGCTCCGTGAAGTTGAGAA	2564
Qy	2418	GAGTCACAAATCTGTAAAAGGCCAAAGCCACAGCCAGAGATTCAGAGGACACACCTAAAAG	2477
Db	2565	GAGTCACAAATCTGTAAAAGGCCAAAGCCACAGCCAGAGATTCAGAGGACACACCTAAAAG	2624
Qy	2478	CAGCTCCAAAAATGCCACCCAGCTCCAACTACCAAGGAAGAGACCACTCCTCCCTCTGC	2537
Db	2625	CAGCTCCAAAAATGCCACCCAGCTCCAACTACCAAGGAAGAGACCACTCCTCCCTCTGC	2684
Qy	2538	CAGCCAAAGCCAGGCCATGGACAAGAAGGCGCAGGGCCCTGGGACCTCCAATTCTGCCAG	2597
Db	2685	CAGCCAAAGCCAGGCCATGGACAAGAAGGCGCAGGGCCCTGGGACCTCCAATTCTGCCAG	2744
Qy	2598	CAGGACCACACTGCCTGCCTCTGGACACCTTCCTATATCTCGGCCCCCTGGAATCGGACC	2657
Db	2745	CAGGACCACACTGCCTGCCTCTGGACACCTTCCTATATCTCGGCCCCCTGGAATCGGACC	2804
Qy	2658	AGATTCTGGCCACGCCCCATCTCAGACTCATCCGTGGAG	2696
Db	2805	AGATTCTGGCCACGCCCCATCTCAGACTCATCCGTGGAG	2843